Análisis y síntesis de los sonidos articulados

Por K. W. WAGNER [1]

534.4

En muchos campos de la Técnica, el conocimiento de la composición exacta de los sonidos que constituyen la voz humana, es de importancia fundamental; pero interesa de modo principalísimo en el dominio de la Telecomunicación, porque es necesario conocer qué relaciones existen entre la inteligibilidad de los sonidos reproducidos y las componentes conservadas del espectro de frecuencias de los sonidos naturales, con el fin de poder proyectar todos los elementos del sistema de comunicación con el ancho de banda estrictamente indispensable para dar paso a las componentes necesarias y suficientes, ya que de ello se derivan consecuencias económicas de valor incalculable en el aprovechamiento óptimo, cada vez más necesario, de los medios de transmisión. Las vocales deben el peculiar carácter que las diferencia de los otros sonidos del lenguaje humano, a ciertos tonos puros, llamados formantes, comprendidos en un dominio de frecuencias dado y cuya altura puede determinarse de un modo bastante exacto. De muy antiguo se ha tratado de analizar estos sonidos vocales mediante la determinación de la situación e intensidad de sus formantes, y de reproducirlos artificialmente valiéndose de diapasones, lengüetas, resonadores y otros órganos mecánicos. Pero el autor fué el primero en utilizar, allá por el año 1 911, otro camino, único que hoy se sigue. Esta investigación le condujo a la preciosa invención de los filtros eléctricos, y con ellos y otros elementos, cuyo carácter es también exclusivamente eléctrico, logró reproducir todas las vocales con absoluta naturalidad. En los azarosos días del derrumbamiento alemán, el autor, que residía en Berlín, perdió todos sus aparatos. En el presente trabajo ofrece a nuestros lectores un interesantísimo resumen de sus investigaciones.

El conocimiento de la composición de los sonidos de la voz humana tiene tanta importancia, desde el punto de vista científico y médico, como por las aplicaciones que de él se derivan a la técnica de la transmisiones, especialmente en telefonía, radio, cinematografía sonora y registro en discos. La primera intervención personal del autor, en esta materia, data de hace cuarenta años, y fué motivada por la necesidad de determinar el ancho de la banda de frecuencias que han de transmitirse en telefonía para lograr una inteligibilidad satisfactoria. Esta cuestión figuraba en el programa de la conferencia de Ingenieros de Telefonía del Estado, que tuvo lugar en Budapest el año 1 911, reunión que fué la precursora del actual Comité Internacional de Telefonía de larga distancia (CCIF), fundado con posterioridad. El Ministerio de Comunicaciones de Berlín encargó al autor la investigación del problema citado, con el fin de informar sobre su resultado en la conferencia de Budapest. Y de estos trabajos nacieron, por una parte, el descubrimiento de los filtros eléctricos, y por otra, los estudios posteriores sobre los sonidos contenidos en la voz humana.

Los filtros eléctricos, que el autor empleó por primera vez, eran entonces todavía muy imperfectos; sus esquemas son los de las *figuras 1, a* y *1, b*; la *figura 1, c* corresponde a un filtro sin distorsión, esto es, que atenúa uniformemente todas las frecuencias. Estos filtros se intercalaban en la línea de transmisión; por a_1 y a_2 se hacían entrar las corrientes telefónicas, para escucharlas en e_1 y e_2. En la rama derivada *bd* hay conectada una inductancia *L*, y como ésta, con frecuencias bajas, tiene poca impedancia, resulta casi un cortocircuito para las mismas, de modo que tales frecuencias sólo llegan a e_1 y e_2 en magnitud insignificante. La red sirve para eliminar las frecuencias bajas, tanto más cuanto más pequeña sea la inductancia. Al comparar la transmisión a través de los filtros representados en *1, a* y *1, c* se pudo conocer el efecto

[1] Doctor en Ciencias Físicas y Electrotécnicas de las Universidades de Gotinga y Darmstadt; Profesor *emeritus* de la Universidad Técnica de Berlín, y antes Director del Instituto "Heinrich Hertz"; Presidente de la Academia de Ciencias y Artes de Maguncia; Miembro de las Academias científicas de Berlín, Boston, Estocolmo y Roma.

de la supresión de las bajas frecuencias en la inteligibilidad de la palabra y, hasta dónde era factible suprimir las citadas frecuencias.

La red *1, b* produce el efecto contrario. El condensador *C* es un cortocircuito para las frecuencias elevadas, de modo que éstas no aparecen prácticamente en e_1 e_2, y tanto menos cuanto más elevado es el valor de *C*. Por consiguiente, la red *1, b* permite determinar hasta qué punto son necesarias las frecuencias elevadas para la inteligibilidad de la conversación.

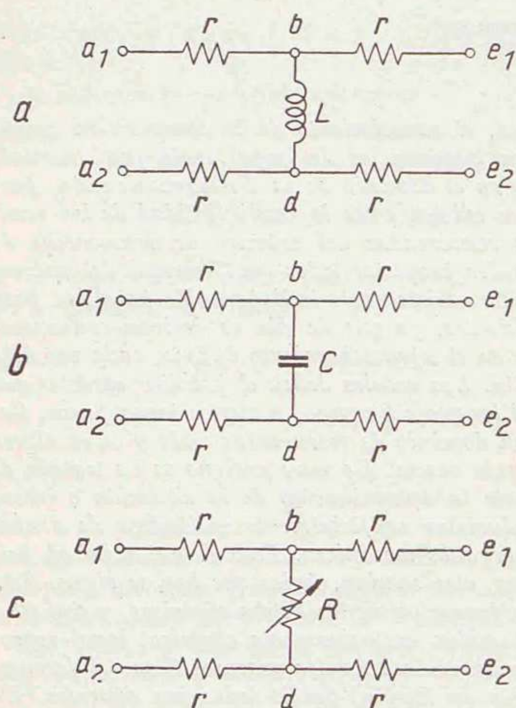


Fig. 1.

El examen teórico de las propiedades de transmisión de redes análogas a las *1, a* y *1, b* dió lugar al invento de los filtros eléctricos, que no son otra cosa que redes que dejan pasar todas las frecuencias comprendidas dentro de ciertos límites y suprimen prácticamente todas las demás. Son, pues, un avance con respecto a las redes *1, a* y *1, b*, ya que éstas tan sólo originan una atenuación, creciente al variar la frecuencia. En los filtros eléctricos, las resistencias puras están reemplazadas por inductancias o capacidades. La *figura 2* muestra algunas redes elementales o células, elegidas entre las más sencillas; la *figura 2, a* corresponde a un filtro de paso bajo, que deja pasar todas las frecuencias inferiores a una dada. Por el contrario, el filtro de la *figura 2, b* solamente da paso a las superiores a una determinada, y por eso se llama filtro de paso alto. El de la *figura 2, c* es un filtro de banda que deja pasar las frecuencias comprendidas dentro de una banda.

Con estos filtros ha sido posible realizar análisis rigurosos de los sonidos contenidos en la voz humana. Anteriormente se habían hecho análisis fonéticos por medio de resonadores acústicos en forma de cavidades y tubos. Estas últimas investigaciones acústicas fueron desarrolladas por el eminente filósofo y psicólogo Carl Stumpf, y su desventaja radica en que son muy laboriosas, requieren mucho tiempo y exigen además oído muy fino y experimentado en el observador. Con el método de los filtros eléctricos, de aplicación mucho más sencilla, se han podido confirmar todos los resultados obtenidos por Stumpf y obtener otros nuevos muy importantes.

Cada sonido vocal es una oscilación periódica, constituída por un tono fundamental y un conjunto de tonos superiores, cuyas frecuencias son múltiplos enteros de la de aquél. La "altura" de un sonido corresponde a la frecuencia del tono fundamental. Ya desde los tiempos de

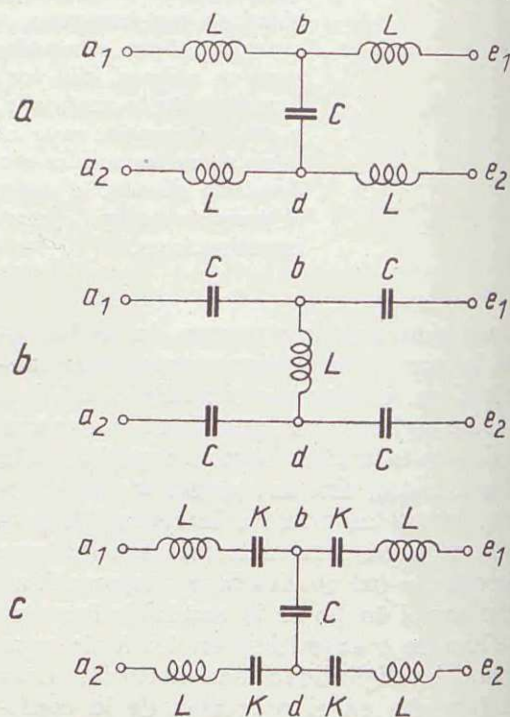


Fig. 2.

Helmholtz se sabe que las fases relativas de los tonos parciales no influyen en el sonido. Se pueden variar arbitrariamente las fases relativas de los diferentes tonos parciales de un sonido sin que el oído note ninguna alteración en el mismo. Esta circunstancia, según veremos más adelante, es importante para la síntesis de las vocales.

Los sonidos producidos por las consonantes son, en cambio, oscilaciones sin un período determinado. Su análisis muestra que están formadas por oscilaciones parciales, cuyas frecuencias cubren una banda sin solución de continuidad.

De lo dicho se infiere que los sonidos de las vocales se pueden representar matemáticamente por series de Fourier, mientras que los de las consonantes han de representarse por integrales de Fourier.

Ya Helmholtz contestó a la pregunta: "¿Qué es lo que caracteriza a un cierto sonido vocal, por ejemplo, el de la *a*?" Helmholtz demostró que cuando se emite una vocal determinada en tonos diferentes, hay un conjunto de vibraciones componentes, comunes a todos ellos, situadas en bandas inmutables de frecuencias. Estas bandas se llaman "formantes" de la vocal respectiva. Cada vocal tiene varias formantes; las que podemos considerar como principales, son características de cada vocal; otras, que llamaremos secundarias, varían de un individuo a otro y dan al sonido su carácter individual. Las diferencias individuales están, pues, determinadas por la intensidad relativa de las formantes secundarias, así como por las diferencias de las correspondientes frecuencias dentro de la banda. En la *tabla I* se indican las frecuencias de las formantes de algunas vocales de la lengua alemana, pronunciadas normalmente. Las formantes principales aparecen subrayadas, y sin subrayar, las secundarias.

TABLA I

u	*o*	*a*	*e*	*i*
350 2 500	450 2 800	700 1 100 1 400 1 800	450 2 600 3 000	300 2 200 3 300

De este cuadro se deduce que la *e* se transforma en *o* si se suprimen las formantes superiores. El mismo procedimiento convierte la *i* en una *u* sorda. Lo cual se puede comprobar fácilmente pronunciando la vocal delante de un micrófono y reproduciéndola en un altavoz. Con un filtro eléctrico intercalado entre el micrófono y el altavoz se puede suprimir cualquier banda de frecuencias.

Para aclarar el concepto de "formante" se muestran en la *figura 3* los resultados de la síntesis de varias vocales, realizada por E. Thienhaus en el Instituto del autor. En coordenadas cartesianas se han llevado frecuencias sobre el eje de abscisas e intensidades en el de ordenadas, utilizando, en este segundo, una escala logarítmica. El tono fundamental tenía, en todos los casos, la frecuencia de 128 Hertz, y de esta frecuencia eran múltiplos enteros los armónicos superiores. Claramente se distinguen las bandas de frecuencias de los armónicos de mayor in-

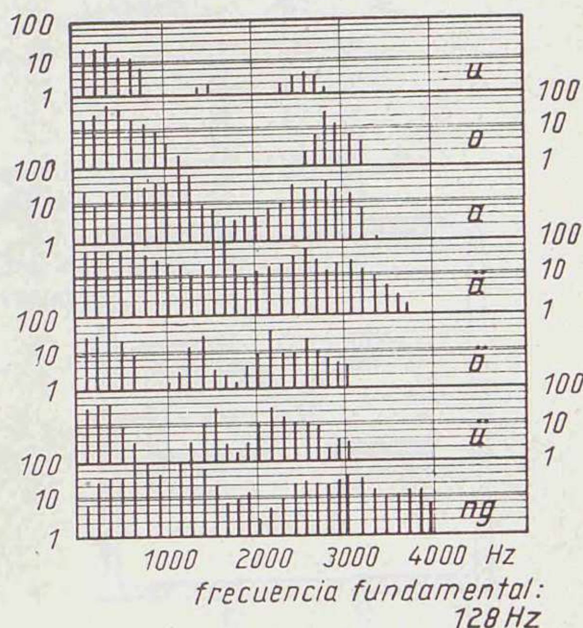


Fig. 3.

tensidad. Estas bandas son las "formantes"; para la *u*, por ejemplo, corresponden a 350 y 2 500 Hertz.

Cuando la misma vocal se pronuncia en otro tono fundamental, varían los armónicos, pero las formantes permanecen invariables. Si se quiere hacer la síntesis de un sonido vocal, es decir, si se desean producir artificialmente vocales, es necesario tener idea clara de cómo se forman naturalmente estas vocales. Partiremos para ello de lo que sabemos sobre la constitución y modo de actuar del órgano humano de la palabra. Dicho órgano está formado, primero, por la laringe, en la que el aire que sale de los pulmones produce oscilaciones, y segundo, por las cavidades anexas, tubo respiratorio, garganta, boca y nariz, que operan como cavidades resonantes o resonadores. Entre todas ellas, es la cavidad bucal la que determina principalmente las características esenciales de los sonidos emitidos, ya que la movilidad de la apertura bucal y de la lengua le confieren una gran posibilidad de variación.

Al comenzar a emitirse un sonido, la laringe

se cierra por la elasticidad de las cuerdas vocales; la presión del aire que sale de los pulmones abre un instante la laringe y ésta se cierra rápidamente al cesar la presión. Este continuo abrir y cerrar produce una sucesión regular de cortos impulsos de aire, cuya frecuencia es igual a la del tono fundamental. Por eso, al variar la ten-

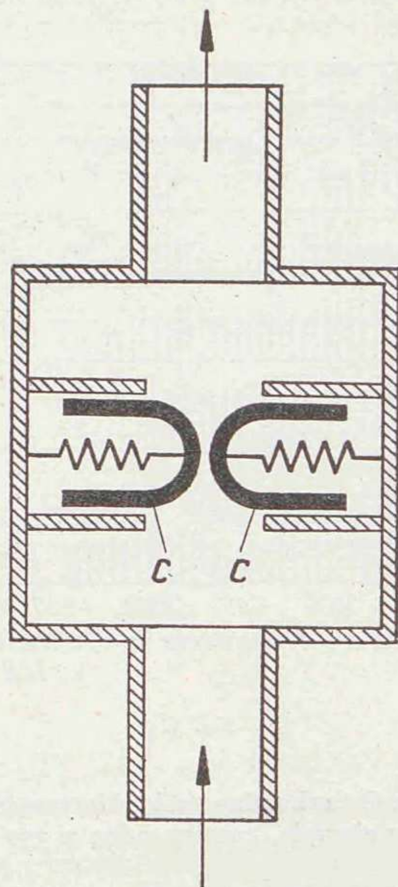


Fig. 4.

sión de las cuerdas vocales, se modifica el tono fundamental, y esto es lo que instintivamente hace la persona que habla o canta cuando quiere modificar la altura del tono.

La *figura 4* muestra un modelo mecánico de laringe, ideado por J. R. Ewald. Las dos almohadillas *c*, mantenidas en contacto por medio de resortes, hacen el papel de las cuerdas vocales, por lo que Ewald designó a este instrumento con el nombre de "Polsterpfeife", cuya traducción literal podría ser "silbato de almohadillas". Si en la dirección que indican las flechas se lanza una corriente de aire sobre las almohadillas, éstas vibran, abriendo y cerrando alternativamente el paso a la corriente y produciendo un ronquido parecido al que emite la laringe aislada del resto del órgano humano de la palabra.

La *figura 5 a* muestra la forma de los imp-sos de aire producidos en la laringe. Para ap-ciar mejor el efecto de tales impulsos sobre sistema de resonadores acoplados entre sí, co-viene descomponer el ruido producido en l armónicos que lo forman, esto es, estudiar espectro de frecuencias, cosa que se ha hecho la parte superior de la *figura 6*.

Si T es el intervalo que media entre dos i-pulsos consecutivos, la frecuencia fundame-tal es:

$$f = \frac{1}{T} .$$

El sonido contiene, además de esta frecuenc todos sus armónicos; la intensidad de todos l de orden inferior a n es prácticamente consta-te, y con respecto a la duración τ del impul existe la relación

$$n f = \frac{1}{\tau} .$$

Por encima de la frecuencia $n f$, la intensid de los armónicos disminuye rápidamente.

Estas relaciones resultan de las consideraci-nes teóricas siguientes:

Si la sucesión de impulsos de la *figura 5 a* s representa por la serie de Fourier

$$\Phi(t) = \sum_n \left[A_n \cos 2 \pi n \frac{t}{T} + \right.$$
$$\left. + B_n \operatorname{sen} 2 \pi n \frac{t}{T} \right] , \qquad [3]$$

cuyos coeficientes se calculan con las fórmul

$$A_n = \frac{2}{T} \int_0^T \Phi(\alpha) \cos 2 \pi n \frac{\alpha}{T} \, d\alpha, \qquad [4a]$$

$$B_n = \frac{2}{T} \int_0^T \Phi(\alpha) \operatorname{sen} 2 \pi n \frac{\alpha}{T} \, d\alpha, \qquad [4b]$$

en que α es la variable de integración. La ampl-tud del armónico de orden n viene dada p

$$C_n = \sqrt{A_n^2 + B_n^2} \qquad [5a]$$

y su fase φ_n resulta de

$$\operatorname{tang} \varphi_n = \frac{A_n}{B_n} . \qquad [5b]$$

Con las expresiones [5*a*] y [5*b*], el armónic de orden n puede escribirse:

$$C_n \operatorname{sen} \left(2 \pi n \frac{t}{T} + \varphi_n \right) . \qquad [6]$$

A veces es útil y recomendable utilizar la expresión compleja,

$$C_n e^{i\varphi_n} = A_n + i\,B_n =$$
$$= \frac{2}{T}\int_0^T \Phi(Z)\,e^{2\pi i n \frac{t}{T}}\,\mathrm{d}Z. \qquad [7]$$

Como el impulso está contenido en el intervalo de tiempo que va desde $t = 0$ hasta $t = \tau$, intervalo muy corto comparado con la duración del período T, en la integral de la ecuación [7], se puede sustituir el límite superior T por τ.

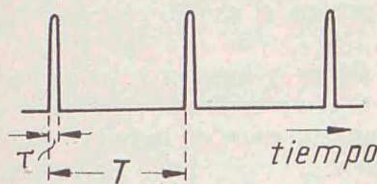


Fig. 5 a.

Puesto que para valores bajos de n, para los que $n \ll \frac{T}{\tau}$, la función exponencial que aparece en dicha integral se puede sustituir por la unidad, se verifica que

$$C_n \simeq \frac{2}{T}\int_0^T \Phi(Z)\,\mathrm{d}Z = \frac{2F}{T}, \qquad [8]$$

en que $F = \int_0^T \Phi(Z)\,\mathrm{d}Z$ es la intensidad del impulso. Se ve, pues, que la *intensidad de los armónicos inferiores es proporcional a la intensidad del impulso, siendo independiente del número de orden* n *y de la forma del impulso.*

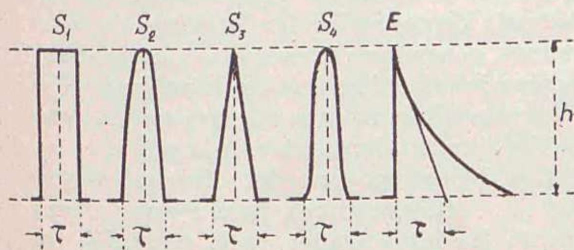


Fig. 5 b.

Para calcular la intensidad de los armónicos superiores, hay que tener en cuenta la forma del impulso. En la *figura 5, b* se ven, a título de ejemplo, algunas formas de impulsos; en ella aparecen un rectángulo, media sinusoide, un triángulo, una sinusoide completa y un impulso exponencial. La letra τ representa la duración del impulso en los cuatro primeros casos, y la constante de tiempo de la exponencial, en el último.

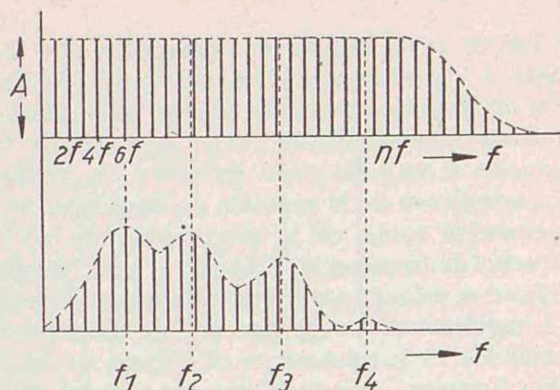


Fig. 6.

A continuación se indican las expresiones que dan las intensidades de los armónicos en los casos anteriores:

$$(S_1)\quad C_n = \frac{2F}{T}\,\frac{\operatorname{sen} x}{x}$$

$$(S_2)\quad C_n = \frac{2F}{T}\,\frac{\cos x}{1-\left(\frac{2x}{\pi}\right)^2}$$

$$(S_3)\quad C_n = \frac{2F}{T}\,\frac{2(1-\cos x)}{x^2}$$

$$(S_4)\quad C_n = \frac{2F}{T}\,\frac{\operatorname{sen} x}{x}\,\frac{\pi^2}{\pi^2 - x^2}$$

$$(E)\quad C_n \simeq \frac{2F}{T}\,\frac{1}{\sqrt{1+\left(2\pi n\frac{\tau}{T}\right)^2}}$$

$$x = \frac{\pi n \tau}{T}.$$

Sin entrar en la discusión de estas fórmulas, nos limitaremos a decir que los valores de C_n, correspondientes a la frecuencia $\frac{1}{\tau}$, para la que comienza la disminución rápida de la intensidad, son:

(S_1) $C_n = 0$
(S_2) $C_n = 0{,}33$
(S_3) $C_n = 0{,}4$
(S_4) $C_n = 0{,}5$
(E) $C_n = 0{,}5$.

Cuando una sucesión de impulsos actúa sobre un sistema de resonadores acoplados, cuyas frecuencias propias

$$f_1, f_2, f_3, f_4 \ldots$$

se hallan en la banda de frecuencias que llega hasta *n f*, y en la que los armónicos de la sucesión de impulsos tienen la misma intensidad, se produce una oscilación, cuyo espectro de frecuencias se ve en la parte inferior de la *figura 6*. Los armónicos de la sucesión de impulsos, cuyas frecuencias están en la proximidad de las frecuencias de resonancia, f_1, f_2, f_3, f_4, . . . , son amplificados, mientras que los restantes se atenúan.

Imaginemos que el sistema de resonadores acoplados es precisamente el órgano de la palabra humana, adaptado para una vocal determinada; en la parte inferior de la *figura 6* vemos el espectro fonético de la vocal respectiva, y f_1, f_2, . . . , serán las formantes correspondientes. Así es que éstas no son otra cosa que las frecuencias de resonancia de las cavidades del órgano de la palabra, adaptadas a la vocal respectiva.

Para poder estudiar mejor los sonidos vocales, el autor construyó un aparato eléctrico que permite reproducirlos artificialmente. Estas investigaciones no sólo interesan al físico, sino que también pueden prestar útiles servicios al especialista en fonética, al lingüista, al psicólogo y al médico.

Ya con anterioridad se habían construído aparatos mecánicos para producir artificialmente las vocales. En un concurso de premios del año 1 779 en la Academia Imperial Rusa, se hicieron estas dos preguntas: "¿Cuáles son el fundamento y las características de los sonidos vocales?" y "¿Será posible construir un aparato capaz de reproducir con naturalidad los sonidos vocales?" Obtuvo el premio el profesor Kratzenstein, que construyó una serie de aparatos, a manera de resonadores, que había que excitar por medio de una corriente de aire, interrumpida por una lengüeta metálica oscilante.

Alrededor del año 1 830, sir David Brewster y E. Willis realizaron experimentos para producir sonidos vocales artificialmente. También usaban resonadores, dotados de lengüetas oscilantes y excitados por una corriente de aire.

F. Wethlo, sobre el mismo principio que el "silbato de almohadillas", de Ewald, construyó modelos de laringe, que insuflaban aire sobre cavidades resonantes, sintonizadas con las frecuencias formantes, con lo que conseguía de este modo sonidos vocales bastante satisfactorios.

R. Paget dió otro paso al construir en plastilina modelos de las cavidades de la boca y garganta, que había que excitar por medio de una laringe artificial. Logró así reproducir todas las vocales de la lengua inglesa.

Todos estos modelos mecánicos tienen la ventaja de que permiten imitar con bastante perfección los órganos naturales de la voz; pero, en cambio, poseen el gran inconveniente de q[illegible] para cada sonido, es necesario un modelo es[illegible] rial, sin ser muy exigentes en cuanto a la na[illegible] ralidad de las voces emitidas. Se compren[illegible] pues, que con tales medios resulte complica[illegible] y laborioso el estudio de la influencia que [illegible] modificaciones de las cavidades del órgano [illegible] la palabra ejercen en el sonido vocal.

Estas consideraciones indujeron al autor a [illegible] construcción de un aparato puramente eléct[illegible] co, con el cual se puede, cómoda y fácilmen[illegible] variar todas las propiedades características [illegible] cada vocal y examinar, de este modo, la [illegible] fluencia que en el sonido ejercen todos los f[illegible] tores.

En la *figura 7* aparece el esquema de dic[illegible] aparato: *O* es un oscilador electromagnético q[illegible] produce una sucesión de impulsos análogos a l[illegible] de la *figura 5*. G, F_1, . . . , F_4, son resonado[illegible] eléctricos, *G* para las oscilaciones fundamental[illegible] y los restantes, para las diferentes formantes. [illegible] hemos visto que para la producción de un to[illegible] fundamental puro y buenas formantes es co[illegible] veniente conectar filtros eléctricos T, S_1, . . . , [illegible] *T* es un filtro de paso bajo, que evita que el r[illegible] sonador *G* deje pasar oscilaciones más altas q[illegible] la fundamental. Los otros son filtros de band[illegible] cuyo paso posee anchura suficiente para u[illegible] buena excitación de oscilaciones en el circui[illegible] formante correspondiente, siempre que se var[illegible] la frecuencia fundamental y se eliminen al mi[illegible] mo tiempo las oscilaciones pertenecientes a l[illegible] bandas de las otras formantes. Los tubos am[illegible] plificadores, tipo 134, situados a la izquierda [illegible] la *figura*, realizan, además de su función ampl[illegible] ficadora, la misión de separar eléctricamente [illegible] entrada de cada circuito formante de todos l[illegible] demás, con lo que se evitan las influencias re[illegible] procas indeseables, que tendrían lugar si se c[illegible] nectaran directamente en paralelo las entrad[illegible] de los circuitos formantes. A la salida de los ci[illegible] cuitos formantes están los divisores de tensi[illegible] A_0, . . . , A_4, que sirven para ajustar adecuad[illegible] mente las intensidades de la oscilación fund[illegible] mental y de las oscilaciones superiores, perten[illegible] cientes a una misma formante. A la salida, tod[illegible] los circuitos formantes están conectados en p[illegible] ralelo, de modo que al combinar las oscilacion[illegible] producidas por ellos, se obtiene la oscilaci[illegible] correspondiente a la vocal deseada. Su intens[illegible] dad se amplifica con los dos tubos amplificad[illegible] res, 134 y 604, hasta que se oiga con suficien[illegible] volumen en el altavoz final. Este altavoz de[illegible] transmitir toda la gama de frecuencias con per[illegible] fecta naturalidad y sin introducir distorsión al[illegible] guna. Finalmente, la sala de audición debe re[illegible] unir buenas condiciones acústicas, ya que cual[illegible]

quier resonancia daría lugar a efectos altamente perturbadores.

Con el equipo descrito, la sintonización de cada formante es completamente independiente de las demás y su manejo resulta en extremo sencillo. Si se quieren imitar los sonidos vocales de determinada persona, conviene comenzar por el análisis de su voz, para obtener indicios o referencias que permitan la adaptación del aparato imitador. Naturalmente, esto se puede llevar a cabo sólo a oído; pero en la práctica sería un procedimiento muy laborioso y de larga duración, ya que el número de variables (amplitud, frecuencia y atenuación de las oscilaciones en cinco circuitos diferentes) admite un número de combinaciones muy grande. En un principio se determinaba oscilográficamente la forma del sonido natural y luego se analizaba el oscilograma, según el método del matemático C. Runge. Este método es laborioso y sólo proporciona los once primeros armónicos, lo que es insuficiente en la mayor parte de los casos. Una dificultad especial radica en que la laringe de la mayor parte de las personas no funciona de modo regular y, por consiguiente, los períodos sucesivos no son iguales, sino solamente semejantes, lo cual obligaba a analizar un cierto número de períodos y hallar la media de los resultados obtenidos. Más cómodo y sencillo es el procedimiento eléctrico, por medio del "tono buscador", que da automáticamente la media.

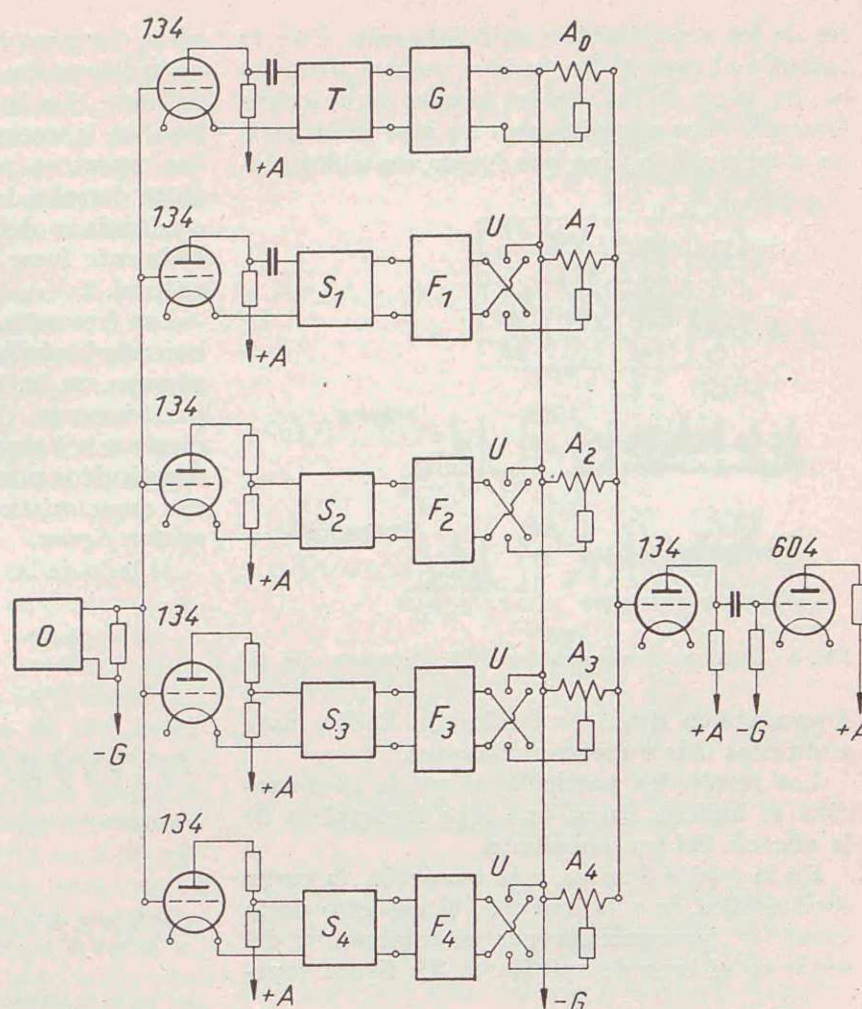


Fig. 7.

Si la persona que se va a examinar tiene una voz musical, el análisis puede llevarse a efecto sin más preparativos. La mayoría de las personas, sin embargo, no puede sostener al hablar la misma altura de tono durante todo el experimento, y por eso los análisis resultan, a menudo, inexactos; pero también esta dificultad se ha podido vencer completamente mediante un procedimiento especial de experimentación.

Las frecuencias formantes, descubiertas por el análisis, se reproducen en el aparato parlante, con lo que se obtiene una vocal bastante parecida a la natural. A continuación se realiza, a oído, el ajuste de las amplitudes y atenuaciones de las diferentes formantes y del tono fundamental.

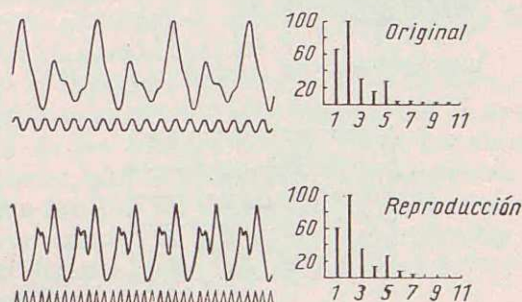


Fig. 8.—*Vocal* u, *pronunciada por "Wz"; frecuencia 156 hz.*

También las frecuencias formantes tienen a veces que ser corregidas. De esta manera se ha llegado a imitar tan perfectamente los sonidos vocales de las personas examinadas, que no se pueden distinguir los sonidos natura-

les de los reproducidos artificialmente. Este es también el caso de los sonidos vocales alemanes ae, oe, ue, y de las vocales nasales de la lengua francesa. Los experimentos no sólo se hicieron en el laboratorio, sino que fueron repetidos pro-

	f Hertz	d	A%
G	211	0,9	–
F_1	662	0,16	100
F_2	1140	0,21	82
F_3	1440	0,06	80

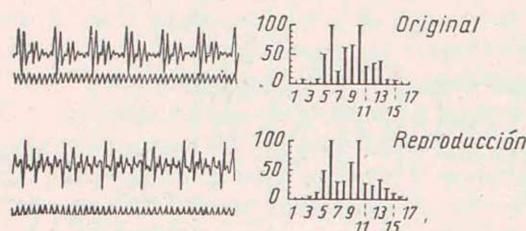
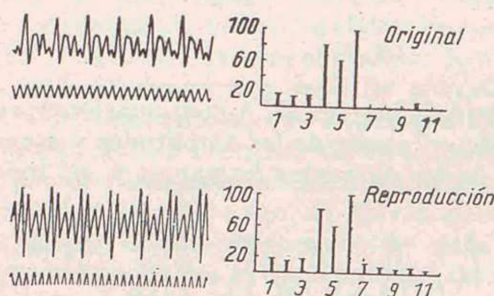


Fig. 9.—*Vocal* a, *pronunciada por "H"; frecuencia 117 hz.*

fusamente en distintos locales de Berlín, ante auditorios más o menos numerosos.

Los resultados particulares que a continuación se indican darán una idea incompleta de la eficacia del procedimiento.

En la *figura 8* se ve, a la izquierda, la curva oscilográfica, y a la derecha, el espectro de la vocal "u", pronunciada por un colaborador del autor en el tono de 156 Hertz. La figura supe-



	f Hertz	d	A%
G	263	1,6	1
F_1	840	0,3	100
F_2	1190	0,1	58
F_3	1900	0,08	6,8

Fig. 10.—*Vocal* a, *pronunciada por "H"; frecuencia 200 hz.*

rior corresponde al sonido natural, y la inferior, al producido artificialmente; ambos sonidos fueron registrados por el mismo micrófono de condensador. La diferencia que se aprecia entre las dos curvas no tiene importancia, ya que la forma depende de las fases relativas de los armó-

nicos componentes, y los valores de dichas fa… no influyen en la sensación percibida a tra… del oído. Lo importante o, mejor aún, lo d… sivo es la concordancia muy aproximada … los espectros parciales, que se muestra en … parte derecha de la misma *figura 8*. A esta co… cordancia se debe que la vocal producida arti… cialmente fuese completamente igual a la vo… natural.

Las *figuras 9 y 10* dan los resultados de … imitaciones de la vocal "a", pronunciada por o… persona, en los tonos de 117 y 200 Hertz, r… pectivamente. En el experimento correspo… diente a la *figura 9*, los circuitos oscilantes est… constituídos por cuatro circuitos formantes, c… yas características se indican en la tabla de … misma *figura*.

Al lado de las formantes intensas, de frecue…

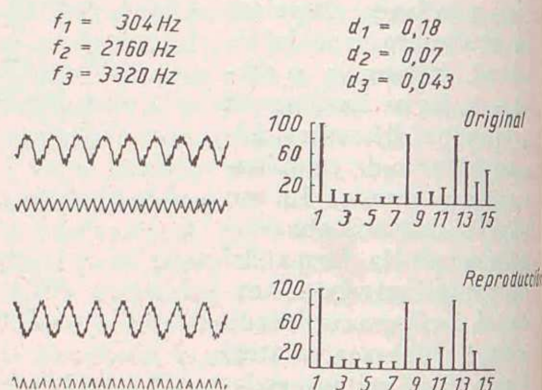


Fig. 11.—*Vocal* i, *pronunciada por "V"; frecuencia 275* …

cias 662 y 1 140 Hertz, aparece otra forman… superior, de intensidad intermedia y 1 440 Her… y una formante inferior, débil, de 211 Hertz.

Si el tono fundamental se eleva a 200 Her… (*figura 10*), baja la formante *G* casi hasta … tono fundamental; F_1 pasa a 840 Hertz, y … permanece casi invariable. La formante d… 1 440 Hertz ya no es perceptible; en cambio, s… presenta una formante secundaria débil en l… 1 900 Hertz. Véanse las magnitudes en la tabl… de la misma *figura 10*.

El desplazamiento parcial de las formant… se debe a que, al variar la altura del tono, el q… habla cambia también, aunque sea muy poc… la posición de la boca. Si se conserva la posici… de la boca y se varía únicamente la altura d… tono por medio de la laringe, no se alteran l… frecuencias de las formantes, aunque se varíe el … tono. Esto se pudo demostrar de modo conclu… yente haciendo que la persona que hablaba pr… nunciase la misma vocal sucesivamente en dif… rentes tonos; si en el aparato se imitaba el s…

nido correspondiente al primer tono, al pasar luego al tono siguiente, no había más que ajustar convenientemente la frecuencia del oscilador para volver a tener una imitación perfecta del sonido natural, siempre que se tuviera cuidado de conservar inalterable la posición de la boca en los diferentes tonos, lo que ciertamente requiere alguna práctica.

En la *figura 11* se ven el oscilograma y el espectro de la vocal "i", correspondientes a los sonidos natural y artificial; el primero, pronunciado en el tono de 275 Hertz. La persona que lo emitía era de voz muy alta y característica, a pesar de lo cual se pudieron imitar sus vocales con toda fidelidad. Las formantes en el caso del sonido de la *figura 11* eran:

$f_1 = 304$ Hz $\qquad d_1 = 0{,}18$
$f_2 = 2\,160$ Hz $\qquad d_2 = 0{,}07$
$f_3 = 3\,320$ Hz $\qquad d_3 = 0{,}043$

La *figura 12* ilustra la imitación de la vocal "o", de la misma persona, en el tono de 167 Hertz. La vocal tiene la formante principal de 350 Hertz ($d = 1{,}0$) y una formante secundaria fuerte en 690 Hertz ($d = 0{,}23$).

En esta persona, el sonido varía muy perceptiblemente con la altura del tono, cosa que ya ha sido explicada, mostrándose ahora los resultados, en la *figura 13*. En un tono bajo, la vocal "o" tiene una formante principal de 400 Hertz, poco más o menos, y una formante secundaria de 600 Hertz (curva formante 1). Si se sube el tono fundamental, la formante secundaria se hace cada vez más débil (curva formante 2) y desparece por completo en el tono fundamental de 200 Hertz (curva formante 3).

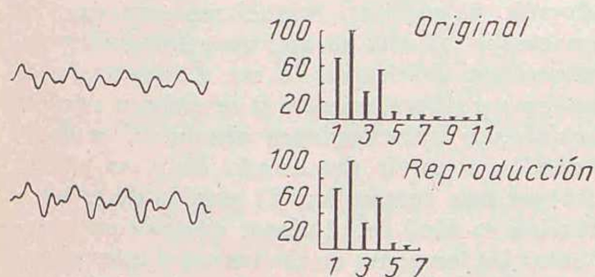


Fig. 12.—*Vocal o, pronunciada por "V"; frecuencia 167 hz.*

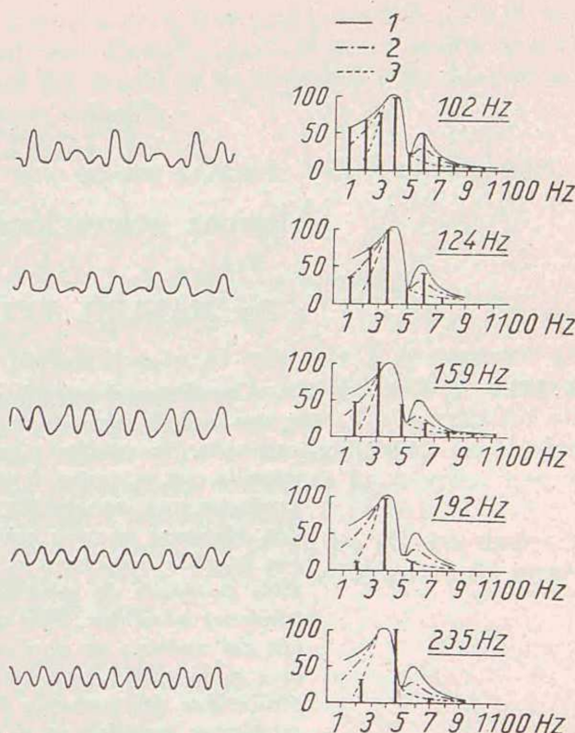


Fig. 13.—*Variación del formador auxiliar con la altura del tono. Vocal o, pronunciada por "V".*

La pregunta, repetidas veces formulada en la literatura técnica, de si las formantes de las vocales son realmente fijas o se alteran con la altura del tono, debe tal vez contestarse diciendo que las formantes son efectivamente inalterables, si el sonido de la vocal no varía. Esto ocurre siempre que la posición del órgano de la palabra no se altere al variar la altura del tono; pero acontece que muchas personas cambian la posición de la boca cuando han de pronunciar espontáneamente cierta vocal con diferentes alturas, y entonces se alteran también las formantes.

Con el aparato descrito se pueden estudiar estas prolijas cuestiones relacionadas con la naturaleza de los sonidos vocales de todas clases, cuestiones que se presentan, por ejemplo, en el estudio fonético de los diferentes dialectos, en la investigación de las alteraciones patológicas de los órganos de la palabra y en otra infinidad de casos.